ESTIMATION OF ASH FUSION TEMPERATURES FROM ELEMENTAL COMPOSITION: A STRATEGY FOR REGRESSOR SELECTION

William G. Lloyd, John T. Riley, Mark A. Risen, Scott R. Gilleland, and Rick L. Tibbitts
Department of Chemistry and Center for Coal Science
Western Kentucky University, Bowling Green, KY 42101

INTRODUCTION

Two developments within the past decade have had a major impact upon our ability to infer useful information about ash fusion temperatures from the composition of the ash. The emergence of fast and accurate multielement analyzers means that the chemical composition of an ash can be determined quickly and reliably. At the same time, the proliferation of personal computers, and of statistical software written for them, makes it possible to rapidly estimate an ash fusion temperature, saving the two or three days' time which would be required to determine fusion temperatures in accordance with standard procedures.[1,2]

In order for such an estimate to be useful, a valid and reliable algorithm is needed. Multiple linear regression (MLR) analysis has been used by a number of workers to obtain estimates generally found to be superior to the single-term factors used in the earlier literature.[3-10] The present work re-examines the use of MLR analysis with particular reference to techniques for avoiding the problem of multicollinearity.

EXPERIMENTAL

Seven source coals, of rank from lignite A to medium volatile bituminous, were selected for study. Coal sources, proximate and ultimate analyses, and ash compositions have been reported.[9,10] Table 1 shows the ranges of the analyses. After reduction to -60 mesh (-0.25 mm) three blends were prepared, in the proportions 3:1, 1:1 and 1:3, for each of the 21 binary combinations of source coals. Ash samples from the source coals and the 63 blended coals were then prepared in accordance with ASTM Method D 1857.[2]

Ash samples were fused with lithium tetraborate for elemental analysis by X-ray fluorescence spectrometry, using an ORTEC Model 6141 spectrometer. Calibration and analysis conditions have been reported elsewhere.[9,10] Cross-analyses were made by inductively coupled plasma spectrometry, using a LECO Plasmarray ICP 500 spectrometer. It is necessary to analyze the ashes of each blend, since composition cannot be estimated by interpolation.[10,11]

Ash fusion temperatures (reducing atmosphere) were measured in duplicate or triplicate on ash splits using a LECO Model AF-600 ash fusibility system. Precision for the four ash fusion temperatures is given in Table 2. Except for fluid temperature, the average error estimate is less than 18°F (10 K). Statistical analyses were conducted using the Statistical Analysis System.[12]

RESULTS

Ash properties cannot be adequately described by assuming a mixture of ten discrete oxides. These components obviously interact with one another, in acid-base and in other metathetic reactions. To take account of these, a set of crossterms can be generated, for example, $[Na_2O]*[SO_3]$ from $[Na_2O]$ and $[SO_3]$.[4,5,7-10] For brevity these

oxides and their crossterms will be represented by the parent element symbols, e.g., Na and Na*S.

Limitations Imposed upon Regressor Selection. In developing an algorithm to estimate an ash fusion temperature, the main task is that of selecting regressor terms from among the ten direct analyses and the 45 crossterms. The wealth of candidate regressors requires the use of some selection rules. There are, for example, 29 million six-term combinations of these regressors. We have chosen to focus upon the extent of collinearity among the selected regressors. When a substantial linear dependence exists between two variables, they are said to be collinear.[13] While high collinearity between a predictive variable and the dependent variable indicates a good predictive relationship, high correlations among predictive variables are not desirable. MLR analysis is based upon the assumption of orthogonality: the distribution of values of each predictor is assumed to be independent of the distribution of values of any other predictor. That ideal condition is seldom found in real data, certainly not in coal ash compositional data. Fortunately, multiple regression can tolerate an appreciable amount of collinearity among regressors. Nevertheless, a major hazard in MLR analysis is that in the presence of excessively high collinearities among regressors (multicollinearity) it is easy to get predictive equations which are good-looking in terms of R^2 and root mean square error of estimate (RMSE), but which are in fact useless.

Multicollinearity in a candidate regression analysis is typically detected by instabilities of regression coefficients, such as[14]

(1) large changes in values when a variable is added to or deleted from the model;
(2) large changes in values when datasets are added or dropped from the model;
(3) large standard errors associated with the coefficients of important terms.

All of these regressions are of the form:

$$AFT_{estimated} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \ldots \qquad (1)$$

The term common to all regressions on a given AFT is the intercept term b_0. A convenient flag for multicollinearity in any particular candidate regression is the standard error of this intercept term (SEI). Since the RMSE's of the better-looking regressions are found to be about 45°F (25 K), we have adopted as a screening criterion that an acceptable regression must have an SEI value of less than 90°F (50 K).

For the following analysis Pearson's correlation coefficient R is used to express collinearity between pairs of predictor variables, while R^2 when used is the square of the multiple correlation coefficient for a given regression.

If we set the critical value of R (R_c) at 0.99, only 5 of the 55 terms (10 direct analyses and 55 crossterms) are found to have collinearities with other regressor terms for which $R > R_c$. The remaining 50 variables are free of 'significant' correlations [significant in terms of this selected value of R_c] with other predictive variables. At $R_c = 0.99$ the five collinear terms occur in a single cluster [Mg, Mg*Na, Mg*S, Mg*Si, Mg*Ti]. On the other hand, at $R_c = 0.85$, only 7 of the 55 terms are free of significant correlations with others, the other 48 occurring in three correlation clusters.

Our first approach has been to set R_c at each of several values, then for each value of R_c to generate the best available regressions which avoid combinations of regressors for which $R_{x,y} > R_c$.

When R_c is set at 0.99 there are few restrictions on regressors. For ash softening temperature we find a good four-term regression, for which R^2 = .859 and RMSE is an acceptable 58.2°F (32.3 K). When the best (by R^2) five-term regressions were examined, the first three regressions were unacceptable, all having SEI's > 200°F. For the first acceptable five-term regression R^2 = 0.885 and RMSE = 52.9°F (29.4 K). Among the next fifty regressions [the best ten (by R^2) regressions containing 6 to 10 terms] none were acceptable; all were rejected by the SEI criterion.

When R_c is set at 0.85 there are considerably more restrictions on the combinations of regressors which can be used in any particular regression. The best (by R^2) three-, four- and five-term regressions are all acceptable by the SEI criterion, but are not as powerful as those found above. For example, for the best five-term regression R^2 = 0.855 and RMSE = 59.3°F (32.9 K). There is an important difference, however, in this second family of regressions: the SEI's for all regressions are well below 90°F, and it is in fact possible to obtain good regressions with ten or more terms present. Table 3 summarizes the fits of the best regressions for these two values of R_c.

It is evident that there are major changes in the extent and complexity of regressor-regressor correlations in this range of R_c. Figures 1 and 2 illustrate the correlations among 19 regressor terms for two intermediate R_c values, 0.98 and 0.90. (Lines connecting terms indicate $R_{XY} > R_c$.) The clustering of the 55 candidate regressor terms as a function of R_c is shown in Table 4.

A Strategy for Selecting Regressors. Based upon preliminary tests, R_c was set at 0.920, and the 19 terms found to be free of correlations were taken as an initial set.

1. The best regressors were selected, by R^2 ranking, from each of the clusters of terms. The largest of these consists of ten terms (Figure 3). To illustrate this selection process with this cluster, the Ca*Ti term is found to make the greatest incremental contribution to the initial set. When this term is selected, the collinearities shown in Figure 3 require that four other terms (Ca, Ca*Na Ca*Si and K) be excluded. With these exclusions, two other terms in this cluster -- Al*Ca and Ca*S -- are isolated from collinearities and are therefore included. Upon analysis of the three remaining terms, Al*K is found to make the greatest incremental contribution to R^2 and is included; and its inclusion requires the rejection of K*Si and K*Ti. A similar R^2 testing procedure was used to select the most useful terms from each of the other clusters.

2. The best 36 regressions (with four, five and six terms) from this enlarged base were then examined, and several regressors -- which appeared in none of the best regressions -- were dropped. Three of the 36 test regressions exhibited multicollinearity. One term, [Si], appeared in all three multicollinear regressions and in none of the 33 good regressions; this term was also dropped.

3. Starting with 22 terms from the above process, steps 1 and 2 were repeated, to ensure that the most useful terms from each cluster were included. After this second iteration, a group of 23 terms remained.

4. To these final terms were added seven additional terms, selected from the pool of remaining terms on the basis of the greatest incremental improvement in overall R^2. For example, if the 24th term is Na*P, overall R^2 is incremented by 0.0063, more than by any other added term; therefore Na*P is added to the set.

Step 4 clearly introduces collinearities. Na*P, for example, is collinear [$R_{x,y} > R_c$] with P*S. Among other added terms Fe*Ti is collinear with Al*Fe, and Fe and Fe*Si are collinear with both Fe*Ti and Al*Fe. The argument for inclusion of the several terms in step 4 is pragmatic: the regressions with these terms are better estimators than those without these terms, and this step still allows overall collinearity to remain at an acceptable level by the SEI criterion.

Best Predictive Regressions. The best regressions obtained under reducing atmosphere for the set of 70 ashes, following the above strategy, are given in Tables 5-8. Calculations have been carried through ten regressor terms. It is possible to generate predictive equations with even more regressors. However, the incremental improvement falls to small values as the number of terms increases. Furthermore, the probability of significance of each regressor term, which typically is > 99.9% for good regressions with as many as nine terms, falls for at least one regressor below 99% with the inclusion of the tenth term. Thus this appears to be a natural break point for these data.

Figures 4-7 show plots of estimated vs. observed fusion temperatures, using the ten-term equations of Tables 5-8. In the tables the average error is estimated using the approximation of average error for large sets:

$$E(avg) = [RMSE]*[2/pi]^{0.5} \qquad (2)$$

The average observed errors of estimate are given in Figures 4-7. These are similar to but slightly lower than the estimated values.

These calculations use data from all 70 ashes. If the three most remote outliers are dropped from each calculation, the average error is decreased by an average of 3.0°F (1.7 K).

Further Testing for Multicollinearity. The most common indicators of regression fit are R^2 and the standard error of fit (RMSE). Table 9 summarizes key characteristics of four regressions on softening temperature, all with good values of R^2 and rmse. On the basis of these indicators alone, the choice would fall between the 30-term and the 55-term regressions. This choice would be unfortunate.

By the SEI criterion [acceptable regressions must have SEI's below 90°F (50 K)] only the first of these four regressions is acceptable, the other three showing SEI's of 800°F and above, indicating excessively high collinearity.

An additional test for multicollinearity is examination of the precision of the regression coefficients. Virtually all regression programs provide an estimate of the standard error associated with each coefficient. We calculate precision as a relative percentage:

$$P (\%) = 100 * [\text{S.E. of coefficient}]/[\text{value of coefficient}] \qquad (3)$$

Coefficients in multiple linear regressions are seldom obtained in high precision, since a moderate displacement in the value of any one coefficient can be balanced by slight shifts in the values of others. For good regressions, precision as defined in Eqn. 3 is typically in the range 5-30%. Table 9 shows the average precision calculated for the ten coefficients of the first regression, and for the first ten coefficients of each of the other regressions. The average error increases tenfold in going to the 20-term regression, and over a hundredfold in going to the 55-term regression.

Another test of the goodness of a regression is made by adding or deleting a dummy variable (a regressor which itself has no predictive power). Instability of a coefficient can then be calculated as:

$$I = [C_{MODIFIED} / C_{ORIGINAL} - 1] * 100\ \% \qquad (4)$$

where $C_{ORIGINAL}$ and $C_{MODIFIED}$ are the regressor coefficients before and after addition/deletion of the dummy variable.

A two-digit random number term was added to each of these regressions, taking the first 70 random numbers listed in a standard statistical reference.[15] For a good regression this dummy variable should have very little effect upon the coefficients of the 'real' regressors. For the 10-term regression in Table 9 the average instability is 0.01%. However, the average instabilities for the first ten terms of the other three regressions are from two to four orders of magnitude larger.

Perhaps the most practical test of the stability of regression coefficients is that of adding or deleting cases from the dataset. If a regression is to have any useful predictive power, it must be reasonably resistant to fluctuation of coefficient values when cases are added or removed. Roughly, variations may be expected to be of the order of magnitude of the precisions of estimate of the coefficients, that is, typically 5-30% for good regressions.

Stability of coefficients to removal of cases was tested by deleting every fifth case in the 70-case dataset, producing a reduced dataset of 56 cases. (This deletion pattern was selected to avoid introduction of systematic bias.) Coefficient instabilities were calculated by Eqn. (4). For the 10-term regression of Table 9 the average instability is 14.1%, consistent with the average coefficient precision of 15.8%. For each of the other regressions in Table 9 the average average instabilities are well over 100%.

These three tests lend support to the use of a critical value of SEI as a convenient indicator of excessive collinearity. The most practical and persuasive showing of the utility of a good regression, however, is to demonstrate its ability to predict from a subset of cases the AFT's of "new" cases. We have taken the coefficients obtained with good ten-term regressions using the reduced set of 56 cases, and have applied them to the 14 excluded cases, treating these as "new" cases. Figure 8 shows the estimated and actual values of softening temperatures for these "new" cases. The average error of estimate is 36.8°F (20.4 K). A similar estimation of hemispherical temperatures yields an average error of estimate of 34.1°F (18.9 K).

DISCUSSION

Gray[7] has recently reviewed various British, American, Australian and international standards for ash fusion temperature determinations. Repeatability (within a laboratory) is 30-40 K for initial deformation temperature, 30 K for hemispherical temperature and 30-50 K for fluid temperature. Tolerated reproducibility (between laboratories) is generally in the range 50-80 K. The repeatability of instrumental ash fusion temperature in this work (Table 2) is considerably tighter than these figures. The average observed error of estimate for the four ash fusion temperatures (Figures 4-7) is 15-18 K. This error includes contributions from coal and ash inhomogeneities, splitting, chemical analysis and AFT determinations, as well as inadequacies of the fitting equations. This approach therefore appears to provide estimates of satisfactory precision.

In the better regressions for estimating softening and hemispherical temperatures certain terms are encountered repeatedly. In softening temperature regressions Al*Ca, Ca*Fe, Na*P and P often occur, always with positive coefficients; Ca*Si, P*S and P*Si also occur frequently, and always with negative coefficients. In hemispherical temperature regressions Al*Ca, Ca*Fe, and Na*P often occur, again always with positive coefficients; Ca*Si, Fe*S and P*S often occur, and always with negative coefficients. Across the ten best ten-term regressions for each AFT the

coefficients in the different regressions are fairly constant, exhibiting standard deviations of 10-20% relative. For example, in ten regressions on hemispherical temperature the coefficients of the common terms and their standard deviations are: Ca*Fe 41,910 +/- 4810, Ca*Si -11,740 +/- 940, Fe*S -47,610 +/- 8300, P*S -411,100 +/- 31,500.

Simple sensitivity analysis calculations have been made to determine, for various regression models, the effect of an analytical error of 1% relative upon the estimated AFT. For the best 10-term regression on softening temperature the average sensitivity for the ten ashes analyzed is 4.0°F per % relative error. The most sensitive analyte is SiO_2, for which a 1% relative analytical error produces an error of estimation of 18°F (10 K). This may be unacceptable for laboratories using atomic absorption analysis, for which repeatability for SiO_2 is 2% absolute or about 5% relative.[16] Using the second best regression (with a loss in RMSE of only 0.1°F) the average sensitivity is 2.9°F per % relative error, and that for SiO_2 is reduced from 18° to 9.8°F. The best ten-term regression on hemispherical temperature shows an average sensitivity of 2.6°F per % relative error, and a sensitivity of 9.2°F per % relative error in SiO_2 determination. The eighth best regression (which gives away 0.4°F in RMSE) shows an average sensitivity of 2.1°F/% and for SiO_2 a sensitivity of 5.2°F/%. As a practical matter it is obviously sensible to determine not only the best valid regression but also a group of regressions, perhaps the ten best valid regressions. The most useful of these can then be selected on the basis of estimated analytical errors and sensitivity analysis for each candidate regression.

ACKNOWLEDGMENT

We gratefully acknowledge financial support of this work through the Robinson Professorship of The Ogden Foundation and a faculty research grant from Western Kentucky University.

REFERENCES

1. Method of Preparing Coal Samples for Analysis, ASTM Method D 2013, Annual Book of ASTM Standards, Vol. 5.05, American Society for Testing and Materials, Philadelphia, PA (published annually).

2. Test Method for Fusibility of Coal and Coke Ash, ASTM Method D 1857 [ISO 540], in reference 1.

3. Rees, O. W., Composition of the Ash of Illinois Coals, Circular 356, Illinois State Geol. Survey, Urbana, IL, 1964.

4. Sondreal, E. A.; Ellman, R. C., Fusibility of Ash from Lignite and its Correlation with Ash Composition, USBM Rept. GFERC/RI-75-1, Pittsburgh, PA, 1975.

5. Winegartner, E. C.; Rhodes, B. T., J. Eng. Power 1975, 97, 395.

6. Vorres, K. W., J. Eng. Power 1979, 101, 497.

7. Gray, V. R., Fuel 1987, 66, 1230.

8. Slegeir, W. A.; Singletary, J. H.; Kohut, J. F., J. Coal Quality 1988, 7, 48.

9. Riley, J. T.; Lloyd, W. G.; Risen, M. A.; Gilleland, S. R.; Tibbitts, R. L., Proc. Intl. Coal Test. Conf. 1989, 7, 58.

10. Lloyd, W. G.; Riley, J. T.; Risen, M. A.; Gilleland, S. R.; Tibbitts, R. L., Energy & Fuels 1990, 4, 360.

11. Riley, J. T.; Gilleland, S. R.; Forsythe, R. F.; Graham, H. D., Jr.; Hayes, F. J., Proc. Intl. Coal Test. Conf. 1989, 7, 32.

12. The Statistical Analysis System, SAS Institute, Inc., Cary, NC 27512.

13. Mason, R. L.; Gunst, R. F.; Hess, J. L., Statistical Design and Analysis of Experiments, Wiley, New York, 1989, p. 490.

14. Chatterjee, S.; Price, B., Regression Analysis by Example, Wiley, New York, 1977, pp 155ff.

15. Arkin, H.; Colton, R. R., Tables for Statisticians, Barnes & Noble, Inc., New York, 2nd ed., 1963, p. 158.

16. Standard Test Method for Major and Minor Elements in Coal and Coke Ash by Atomic Absorption, ASTM Method D 3682, in reference 1.

Table 1

Ranges of Properties of Coals and Ashes[a]

	Low	Mean	High
Coals			
rank	ligA	---	mvb
moisture	1.1%	10.9%	18.4%
% ash[b]	5.8%	11.4%	18.9%
% volatile matter[b]	25.3%	38.2%	44.1%
% carbon[b]	55.3%	67.5%	84.4%
% hydrogen[b]	3.33%	4.19%	5.06%
% nitrogen[b]	0.38%	1.04%	1.57%
% sulfur[b]	0.34%	1.08%	3.78%
Btu/lb[b]	8,770	11,500	14,700
MJ/kg[b]	20.4	26.7	34.1
Hardgrove Index	34	47	90
Free Swelling Index	0	2	9
Ash Composition			
SiO_2	33.1%	40.1%	50.4%
Al_2O_3	14.7%	22.2%	24.7%
CaO	0.55%	10.6%	26.9%
Fe_2O_3	3.49%	8.8%	24.1%
SO_3	0.72%	7.8%	10.5%
MgO	1.06%	2.28%	4.59%
K_2O	0.16%	1.37%	2.63%
P_2O_5	0.01%	0.77%	2.36%
TiO_2	0.82%	1.11%	1.66%
Na_2O	0.51%	0.91%	1.20%

[a] Coal data are from the seven source coals; ash data are from the suite of 70 ashes. [b] Calculated on a dry, ash-included basis.

Table 2

Precision of Ash Fusion Temperature Determinations[a]

Fusion Temperature, °F	DF	Std dev.°F	Avg Error,°F[b]
Initial Deformation	94	20.2	16.1
	(89)	(16.1)	(12.9)
Softening	95	13.4	10.7
	(91)	(11.1)	(8.9)
Hemispherical	94	21.8	17.4
	(89)	(19.1)	(15.3)
Fluid	93	42.0	33.5
	(89)	(35.5)	(28.3)

[a] First figures given are obtained with all 70 ashes. Figures in parentheses are obtained after rejection of three outlier ashes.

[b] Estimated by Eqn. (2).

Table 3

Best Regression Fits at Two Values of R(critical)[a]

	R(critical) = 0.99		R(critical) = 0.85	
terms	R^2	RMSE, °F	R^2	RMSE, °F
3	.825	64.2°	.766	74.2°
4	.859	58.2	.804	68.5
5	.885	52.9[b]	.855	59.3
6	---	[c]	.879	54.5
7	---	[c]	.895	51.2
8	---	[c]	.908	48.5
9	---	[c]	.914	47.2
10	---	[c]	.921	45.5

[a] Regressions meeting the criterion that SEI < 90°F [< 50 K].

[b] Fourth best regression ranked by R^2; SEI's of first three regressions > 200°F.

[c] None of the first ten regressions ranked by R^2 have SEI's below 90°F.

Table 4

Correlation Clusters of Regressors at Various Levels of $R_{critical}$

$R_{critical}$	terms in clusters	total terms in clusters	total free terms
0.99	5	5	50
0.98	6, 4, 3, 2, 2, 2	19	36
0.95	7, 6, 5, 4, 4, 3, 2	31	24
0.92	10, 7, 7, 5, 5	36	19
0.90	19, 8, 7, 5	39	16
0.88	27, 7, 5	41	14
0.85	40, 5, 3	48	7

Table 5

Best Regressions on Initial Deformation Temperature

Intcpt.°F	Terms and Coefficients					SEI,°F	R^2	RMSE,°F	average error,°F
	S*Si	K*P	Ca*Fe	Al*Na					
1,863°F	-8.01E3	1.26E6	1.61E4	1.40E5		36.4	.803	72.1	57.5
	Na*Si	Fe*S	Ca*Ti	Ca*Fe	Al*Na				
1,822	5.60E4	-5.08E4	-3.49E5	7.85E4	1.43E5	34.3	.844	64.8	51.7
	P	S*Si	Ca*Ti	Ca*Fe	Fe*Na				
2,093	1.66E4	-1.82E4	-2.50E5	6.26E4	-3.99E5				
	Al*Na								
	2.96E5					37.1	.877	58.0	46.3
	P	Mg*P	S*Si	Ca*Ti	Ca*Fe				
2,138	1.37E4	1.09E5	-1.94E4	-2.87E5	6.63E4				
	Fe*Na	Al*Na							
	-4.54E5	3.20E5				41.1	.886	56.1	44.8
	P	S*Si	P*Si	P*S	Ca*Ti				
2,080	1.46E5	-1.81E4	-2.49E5	-3.50E5	-3.23E5				
	Ca*Fe	Fe*Na	Al*Na						
	6.28E4	-4.43E5	3.50E5			50.4	.9058	51.5	41.1
	P	P*Si	K*S	P*S	Fe*K				
1,654	1.95E5	-3.08E5	-1.55E5	-6.17E5	1.21E5				
	Ca*Ti	Ca*Fe	Fe*Na	Al*Na					
	-3.90E5	7.68E4	-7.73E5	3.71E5		68.8	.9147	49.5	39.5
	P	P*Si	K*S	P*S	K*Ti				
1,609	1.94E5	-3.21E5	-1.63E5	-5.82E5	7.69E5				
	Fe*K	Ca*Ti	Ca*Fe	Fe*Na	Al*Na				
	9.03E4	-3.54E5	7.40E4	-7.30E5	3.50E5	69.8	.9212	47.9	38.2

Table 6

Best Regressions on Softening Temperature

Intcpt.°F	Terms and Coefficients					SEI,°F	R^2	RMSE,°F	average error,°F
	P*S	K*P	Na*P	Al*Na					
2,004°F	-3.76E5	6.92E5	3.15E6	7.30E4		27.5	.852	59.5	47.5
	Ca*Si	P*S	Al*K	Al*Ca	Na*P				
1,896	-5.34E3	-5.52E5	4.09E4	1.48E4	5.64E6	54.4	.882	53.5	42.7
	Ca*Si	Fe*Si	K*S	K*Na	Ca*Fe				
2,344	-9.04E3	-5.86E3	-5.50E5	7.07E6	5.80E4				
	Fe*Na								
	-3.62E5					55.8	.9047	48.5	38.7
	P	Ca*Si	P*Si	P*S	Al*K				
1,827	1.23E5	-9.31E3	-2.20E5	-8.26E5	5.21E4				
	Al*Ca	Na*P							
	2.23E4	4.64E6				47.9	.9240	43.6	34.8
	Ca*Si	Fe*Si	P*S	Fe*S	K*Na				
2,212	-1.09E4	-2.73E3	-4.27E5	-4.33E4	1.69E6				
	Ca*Fe	Al*Ca	Na*P						
	5.15E4	8.48E3	4.20E6			65.3	.9286	42.7	34.0
	Ti	Fe*Mg	Si*Ti	K*S	P*S				
1,806	1.33E5	-1.75E4	-2.89E5	-3.03E5	-5.48E5				
	K*Na	Fe*K	Al*K	Na*P					
	2.62E6	-7.48E4	1.10E5	4.64E6		53.5	.9356	40.9	32.6
	Fe	Mg*P	Fe*Mg	Al*Fe	S				
2,248	2.97E3	2.29E5	-8.57E4	-2.67E4	1.37E4				
	S*Si	Al*K	Ca*Na	Ca*Fe	Al*Na				
	-5.80E4	7.69E4	-2.28E5	4.65E4	3.28E5	46.8	.9420	39.1	31.2

Table 7

Best Regressions on Hemispherical Temperature

Intcpt, °F	Terms and Coefficients					SEI, °F	R^2	RMSE, °F	average error, °F
	P*S	K*P	Na*P	Al*Na					
2,078	-3.53E5	7.93E5	2.77E6	6.21E4		25.2	.872	54.6	43.6
	Ca*Si	P*S	K*Ti	Al*Ca	Na*P				
1,897	-5.35E3	-4.80E5	1.09E6	1.61E4	4.75E6	69.3	.896	49.7	39.6
	Ca*Si	P*S	Fe*S	K*P	Ca*Fe				
2,070	-4.41E3	-1.09E5	-3.62E4	1.29E6	4.14E4				
	Al*Na								
	1.18E5					25.8	.9127	45.9	36.6
	Ca*Si	K*S	P*S	Fe*S	Ca*Fe				
2,195	-1.18E4	1.34E5	-5.72E5	-4.28E4	4.01E4				
	Al*Ca	Na*P							
	1.46E4	5.60E6				23.1	.9287	41.8	33.4
	P	Ca*Si	P*S	Fe*S	K*Na				
2,066	1.81E4	-1.16E4	-3.61E5	-3.04E4	1.14E6				
	Ca*Fe	Al*Ca	Na*P						
	3.42E4	1.73E4	2.22E6			36.7	.9379	39.3	31.4
	Fe	Ca*Si	Fe*Si	P*S	Fe*S				
2,192	6.54E3	-1.11E4	-1.62E4	-4.65E5	-4.26E4				
	Al*K	Ca*Fe	Al*Ca	Na*P					
	5.33E4	3.34E4	1.43E4	4.55E6		66.2	.9446	37.5	29.9
	P	Fe	Ca*Si	Fe*Si	P*S				
2,256	9.71E5	4.53E3	-1.21E4	-1.16E4	-3.88E5				
	Fe*S	K*Na	Ca*Fe	Al*Ca	Na*P				
	-4.82E4	1.40E6	4.31E4	1.22E4	2.98E6	74.6	.9477	36.7	29.3

Table 8

Best Regressions on Fluid Temperature

Intcpt, °F	Terms and Coefficients					SEI, °F	R^2	RMSE, °F	average error, °F
	S*Si	P*S	K*P	Al*Na					
2,185°F	-8.74E3	-6.09E4	1.00E6	1.73E5		24.9	.892	53.1°F	42.3°F
	Mg*P	S*Si	P*S	K*P	Al*Na				
2,171	9.97E4	-8.52E3	-1.06E5	1.06E6	1.78E5	24.6	.9010	51.2	40.8
	S*Si	K*P	Ca*Ti	Ca*Fe	Fe*Na				
2,319	-1.88E4	6.33E5	-1.11E5	3.30E4	-2.48E5				
	Al*Na								
	2.90E5					34.0	.9172	47.1	37.6
	Mg*P	Fe*Mg	S*Si	P*S	K*P				
2,155	2.19E5	-3.45E4	-1.29E4	-1.39E5	9.50E5				
	Ca*Fe	Al*Na							
	1.27E4	2.40E5				23.2	.9266	44.8	35.7
	Ti	Mg*P	Fe*Mg	S*Si	Si*Ti				
2,203	6.66E4	2.33E5	-3.91E4	-1.45E4	-1.40E5				
	P*S	K*P	Al*Na						
	-2.05E5	1.10E6	2.44E5			55.3	.9313	43.7	34.8
	Ti	Mg*P	Fe*Mg	S*Si	Si*Ti				
2,204	5.02E4	2.35E5	-4.41E4	-1.57E4	-1.10E5				
	P*S	K*P	Ca*Fe	Al*Na					
	-1.89E5	1.11E6	7.83E3	2.57E5		54.1	.9355	42.7	34.0
	Ti	Mg*P	Fe*Mg	S*Si	P*Si				
2,283	6.36E4	2.35E5	-5.13E4	-1.15E4	5.66E4				
	Si*Ti	P*S	K*Na	K*Ti	Al*Na				
	-2.28E5	-2.81E5	-2.01E6	2.21E6	3.14E5	79.9	.9391	41.8	33.3

Table 9

Indicators of Multicollinearity in Softening Temperature Regressions

	10 terms[a]	20 terms[b]	30 terms[b]	55 terms
R^2 (uncorrected)	.942	.9547	.9667	.9770
rmse	39.1°F 21.7 K	37.9°F 21.1 K	36.4°F 20.2 K	50.5°F 28.1 K
SEI	46.8°F 26.0 K	813°F 452 K	>2120°F >1170 K	>21,000°F >11,600 K
Avg precision of first 10 coefficients[c]	15.8 %	159 %	272 %	>2,400 %
Avg instability of first 10 coefficients to a dummy variable[d]	0.01%	4.6 %	32.1%	400 %
Avg instability of first 10 coefficients to removal of 14 cases[e]	14.1 %	111 %	449 %	>1,500 %

[a] Best regression from Table 6.

[b] Selected by forward selection procedure.

[c] [S.E. of coefficient]/[coefficient] * 100%.

[d] Average shift in value of coefficients upon introduction of a random number variable (see text).

[e] Average shift in value of coefficients upon deletion of every fifth case in the dataset.

FIGURE 1

CORRELATIONS AMONG 19 TERMS

R(crit) = 0.98

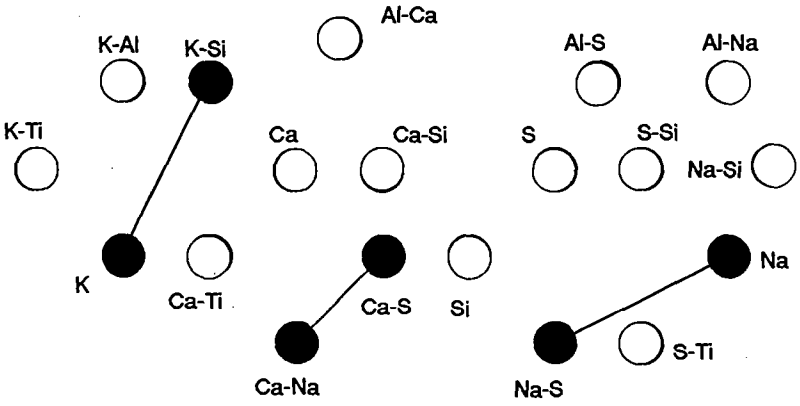

FIGURE 2

CORRELATIONS AMONG 19 TERMS

R(crit) = 0.90

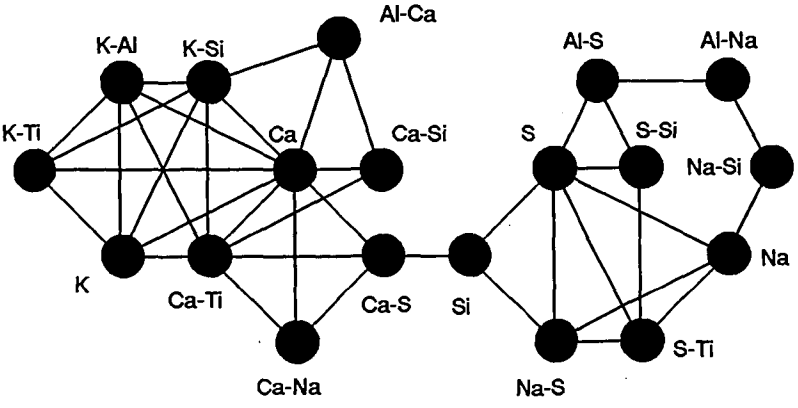

FIGURE 3

SELECTION TO AVOID COLLINEARITIES

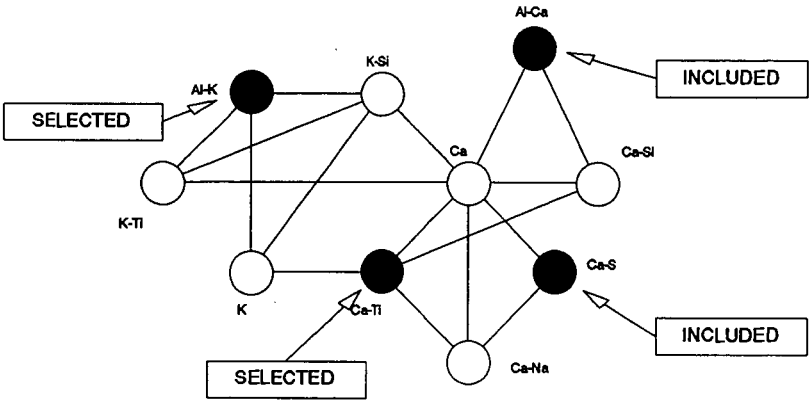

FIGURE 4

ESTIMATE OF INITIAL DEFORMATION TEMP.
70 CASES

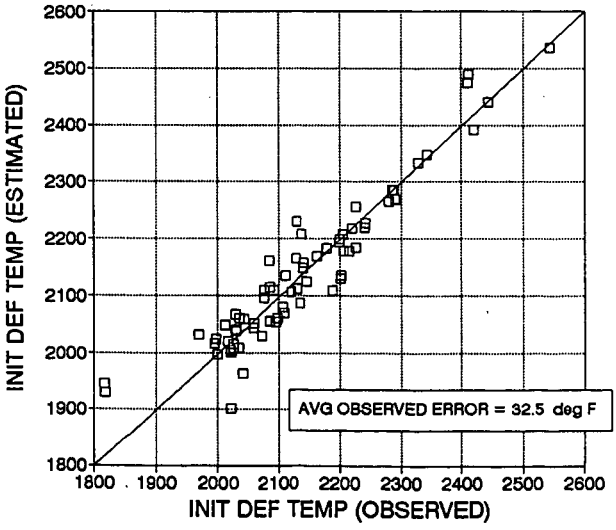

FIGURE 5

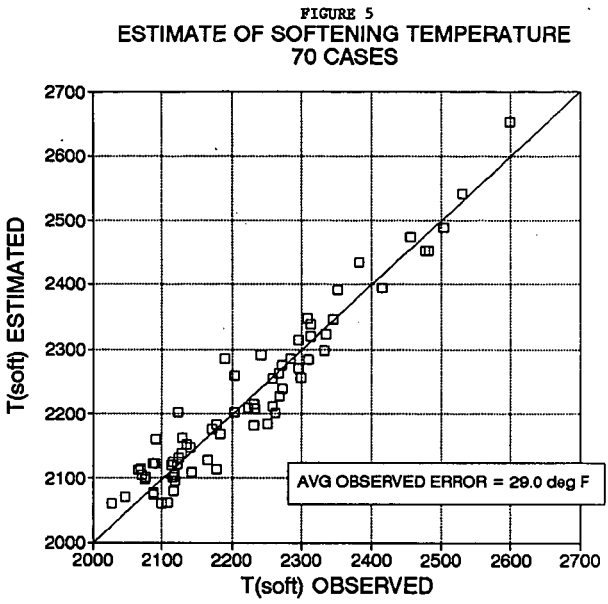

FIGURE 6

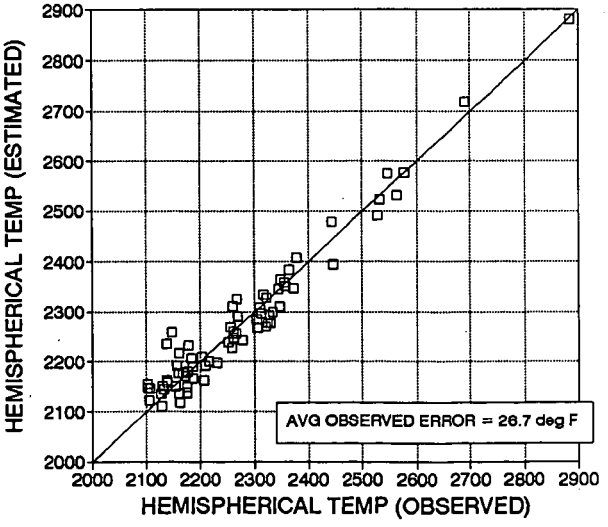

FIGURE 7

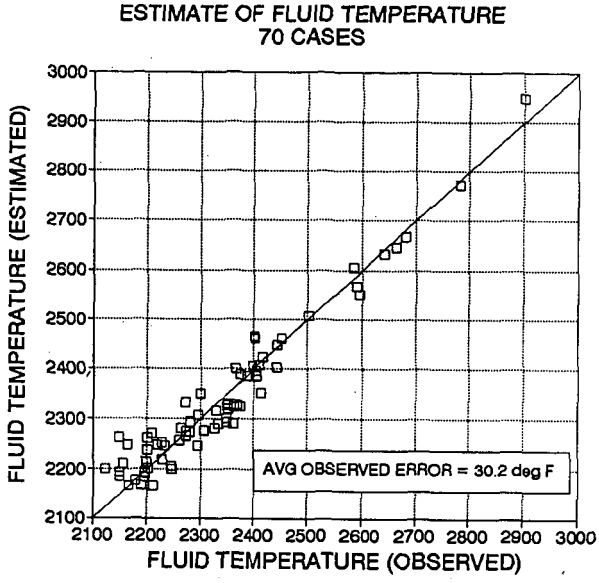

FIGURE 8